Are the Digits of Pi Random?

David H. Bailey
Lawrence Berkeley National Laboratory
Berkeley, CA, USA 94720
dhbailey@lbl.gov

Richard E. Crandall
Reed College
Portland, OR 97202
crandall@reed.edu

The First 1000 Decimal Digits of Pi

3.

1415926535897932384626433832795028841971693993751058209749445923078164062862089986280348253421170679
8214808651328230664709384460955058223172535940812848111745028410270193852110555964462294895493038196
4428810975665933446128475648233786783165271201909145648566923460348610454326648213393607260249141273
7245870066063155881748815209209628292540917153643678925903600113305305488204665213841469519415116094
3305727036575959195309218611738193261179310511854807446237996274956735188575272489122793818301194912
9833673362440656643086021394946395224737190702179860943702770539217176293176752384674818467669405132
0005681271452635608277857713427577896091736371787214684409012249534301465495853710507922796892589235
4201995611212902196086403441815981362977477130996051870721134999999837297804995105973173281609631859
5024459455346908302642522308253344685035261931188171010003137838752886587533208381420617177669147303
5982534904287554687311595628638823537875937519577818577805321712268066130019278766111959092164201989

There is at least one unusual feature in these digits. Can you find it?

A Brief History of Pi

2000 BCE	Babylonians	$\pi = 3.125.$
550 BCE	Hebrews (1 Kings 7:23)	$\pi = 3$
250 BCE	Archimedes	$\pi \approx 3.1418$
150	Ptolemy	$\pi \approx 3.14166$
480	Tsu Ch'ung Chi	$\pi \approx 3.1415926$
1593	Viete	$\pi \approx 3.1415926536$
1665	Newton	π to 16 decimal places
1706	Machin	π to 100 decimal places
1767	Lambert and Legendre	Proved π is irrational
1874	Shanks	π to 527 decimal places
1882	Lindemann	Proved π is transcendental
1961	Shanks and Wrench	π to 100,000 decimal places
1973	Guilloud and Bouyer	π to one million decimal places
1976	Brent and Salamin	Quadratically convergent algorithm for π
1986	Borweins	Quartically (fourth order) convergent algorithm for π
1989	Chudnovskys	π to one billion decimal places
1997	3 authors	Algorithm for computing n-th hexadecimal digit of π
1997	3 authors	Hexadecimal digits of π starting at 10 billionth digit
1999	Kanada and Tamura	π to 206 billion decimal places
1999	Percival	Hexadecimal digits of π starting at 1.25 trillionth digit

Normality

The real number α is *normal* to base b if every sequence of k consecutive digits in the base-b expansion of α appears with limiting frequency b^{-k}. We say that α is *absolutely normal* if α is normal to every integer base $b \geq 2$.

Widely believed to be absolutely normal:

- π and e
- $\log 2$ and $\sqrt{2}$
- the golden mean $\tau = (1+\sqrt{5})/2$
- *every* irrational algebraic number
- many other "natural" irrational constants

But there are *no* proofs — not for any of these constants, not for any base. Even a weaker "digit-dense" property has not been established for any of these constants. Normality proofs exist only for artifically constructed constants such as 0.1234567891011121314...

Two Questions

1. Let $x_0 = 0$, and

$$x_n = \left(2x_{n-1} + \frac{1}{n}\right) \bmod 1$$

Is (x_n) equidistributed in $[0, 1)$?

2. Let $x_0 = 0$ and

$$x_n = \left(16x_{n-1} + \frac{120n^2 - 89n + 16}{512n^4 - 1024n^3 + 712n^2 - 206n + 21}\right) \bmod 1$$

Is (x_n) equidistributed in $[0, 1)$?

Consequences

If answer to Question 1 is "yes", then $\log 2$ is normal to base 2.

If answer to Question 2 is "yes", then π is normal to base 16 (and hence to base 2 also).

Hypothesis A

Denote by $r_n = p(n)/q(n)$ a rational-polynomial function, $0 \leq \deg(p) < \deg(q)$. Let b be an integer, $b \geq 2$ and set $x_0 = 0$. Then the sequence

$$x_n = (bx_{n-1} + r_n) \bmod 1$$

either has a finite attractor or is equidistributed in $[0, 1)$.

Theorem 1: Assuming Hypothesis A, each of the constants π, $\log 2$, $\zeta(3)$ is normal to base 2. Also, on Hypothesis A, if $\zeta(5)$ is irrational then it likewise is normal to base 2.

This list of constants is merely representative — numerous other constants could also be listed here.

Background: A New Formula for Pi

This formula was found in 1997 by a computer program, using the PSLQ integer relation algorithm:

$$\pi = \sum_{k=0}^{\infty} \frac{1}{16^k} \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right)$$

This formula may be used to compute the n-th hexadecimal (or binary) digit of π, without computing any of the first $n-1$ digits.

Here is a formula of this same type for $\log 2$:

$$\log 2 = \sum_{k=1}^{\infty} \frac{1}{k 2^k}$$

Although this formula has been known for centuries, the connection to computing individual binary digits of $\log 2$ was only very recently discovered.

The BBP Algorithm for Computing Individual Hex Digits of Pi

Let S_1 be the first of the four sums in the formula for π.

$$
\begin{aligned}
(16^n S_1) \bmod 1 &= \left(\sum_{k=0}^{\infty} \frac{16^{n-k}}{8k+1} \right) \bmod 1 = \left(\sum_{k=0}^{n} \frac{16^{n-k}}{8k+1} + \sum_{k=n+1}^{\infty} \frac{16^{n-k}}{8k+1} \right) \bmod 1 \\
&= \left(\sum_{k=0}^{n} \frac{16^{n-k} \bmod 8k+1}{8k+1} + \sum_{k=n+1}^{\infty} \frac{16^{n-k}}{8k+1} \right) \bmod 1
\end{aligned}
$$

1. Compute each numerator of each term in the first sum using the binary algorithm for exponentiation, reducing each product modulo $8k+1$.
2. Divide each numerator by its respective $8k+1$.
3. Sum the terms of the first series, discarding integer parts.
4. Compute the second sum (just a few terms are needed).
5. Add the two sum results, again discarding the integer part.
6. Repeat for S_1, S_2, S_3, S_4, and calculate $4S_1 - 2S_2 - S_3 - S_4$.
7. The resulting fraction, when expressed in hexadecimal format, gives the first few hex digits of π beginning at position $n+1$.

Ordinary 64-bit or 128-bit floating-point arithmetic suffices for these operations — multiple precision arithmetic software is *not* required.

Some Computational Results

Position	Hex Digits of π Starting at Position
10^6	26C65E52CB4593
10^7	17AF5863EFED8D
10^8	ECB840E21926EC
10^9	85895585A0428B
10^{10}	921C73C6838FB2
10^{11}	9C381872D27596
1.25×10^{12}	07E45733CC790B

Thanks to Fabrice Bellard of France and Colin Percival of Canada.

Some Other Constants with Base 2 BBP-Type Formulas

$$
\begin{aligned}
\log 3 &= \sum_{k=0}^{\infty} \frac{1}{4^k(2k+1)} \\
\log 7 &= \frac{3}{4}\sum_{k=0}^{\infty} \frac{1}{8^k}\left(\frac{2}{8k+1} + \frac{1}{8k+2}\right) \\
\pi^2 &= \frac{1}{8}\sum_{k=0}^{\infty} \frac{1}{64^k}\left(\frac{144}{(6k+1)^2} - \frac{216}{(6k+2)^2} - \frac{72}{(6k+3)^2} - \frac{54}{(6k+4)^2} + \frac{9}{(6k+5)^2}\right) \\
\log^2 2 &= \frac{1}{6}\sum_{k=0}^{\infty} \frac{1}{16^k}\left(\frac{16}{(8k+1)^2} - \frac{40}{(8k+2)^2} - \frac{8}{(8k+3)^2} - \frac{28}{(8k+4)^2}\right. \\
&\quad \left. - \frac{4}{(8k+5)^2} - \frac{10}{(8k+6)^2} + \frac{2}{(8k+7)^2} - \frac{3}{(8k+8)^2}\right) \\
\pi^2 - 6\log^2 2 &= 12\sum_{k=1}^{\infty} \frac{1}{k^2 2^k} \\
\pi\sqrt{3} &= \frac{9}{32}\sum_{k=0}^{\infty} \frac{1}{64^k}\left(\frac{16}{6k+1} - \frac{8}{6k+2} - \frac{2}{6k+4} - \frac{1}{6k+5}\right)
\end{aligned}
$$

A Base 2 BBP-Type Formula for $\zeta(3)$

$$\zeta(3) = \sum_{k=1}^{\infty} \frac{1}{4096^k} \frac{p(k)}{q(k)}$$

where

$$\begin{aligned}
\frac{7}{8}\frac{p(m)}{q(m)} = {} & \frac{1}{2\,(1+8\,m)^3} + \frac{1}{4\,(2+8\,m)^3} - \frac{1}{16\,(3+8\,m)^3} - \frac{1}{16\,(4+8\,m)^3} \\
& - \frac{1}{128\,(5+8\,m)^3} + \frac{1}{256\,(6+8\,m)^3} + \frac{1}{1024\,(7+8\,m)^3} + \frac{3}{(1+24\,m)^3} \\
& - \frac{21}{(2+24\,m)^3} - \frac{3}{2\,(3+24\,m)^3} + \frac{15}{(4+24\,m)^3} - \frac{3}{4\,(5+24\,m)^3} \\
& - \frac{21}{4\,(6+24\,m)^3} + \frac{3}{8\,(7+24\,m)^3} + \frac{3}{16\,(9+24\,m)^3} - \frac{21}{16\,(10+24\,m)^3} \\
& - \frac{3}{32\,(11+24\,m)^3} + \frac{15}{16\,(12+24\,m)^3} - \frac{3}{64\,(13+24\,m)^3} \\
& - \frac{21}{64\,(14+24\,m)^3} + \frac{3}{128\,(15+24\,m)^3} + \frac{3}{256\,(17+24\,m)^3} \\
& - \frac{21}{256\,(18+24\,m)^3} - \frac{3}{512\,(19+24\,m)^3} + \frac{15}{256\,(20+24\,m)^3} \\
& - \frac{3}{1024\,(21+24\,m)^3} - \frac{21}{1024\,(22+24\,m)^3} + \frac{3}{2048\,(23+24\,m)^3}.
\end{aligned}$$

A similar, but even more complicated, formula exists for $\zeta(5)$.

Some Base 3 BBP-Type Formulas

$$
\begin{aligned}
\log 2 &= \frac{2}{27} \sum_{k=0}^{\infty} \frac{1}{81^k} \left(\frac{9}{4k+1} + \frac{1}{4k+3} \right) \\
&= \sum_{n=0}^{\infty} \frac{1}{9^n (2n-1)} \\
\pi^2 &= \frac{2}{27} \sum_{k=0}^{\infty} \left(\frac{243}{(12k+1)^2} - \frac{405}{(12k+2)^2} - \frac{81}{(12k+4)^2} - \frac{27}{(12k+5)^2} \right. \\
&\quad \left. - \frac{72}{(12k+6)^2} - \frac{9}{(12k+7)^2} - \frac{9}{(12k+8)^2} - \frac{5}{(12k+10)^2} + \frac{1}{(12k+11)^2} \right) \\
6\sqrt{3} \tan^{-1} \left(\frac{\sqrt{3}}{7} \right) &= \sum_{k=0}^{\infty} \frac{1}{27^k} \left(\frac{3}{3k+1} + \frac{1}{3k+2} \right)
\end{aligned}
$$

A Base 5 BBP-Type Formula

$$
\frac{25}{2} \log \left(\frac{781}{256} \left(\frac{57 - 5\sqrt{5}}{57 + 5\sqrt{5}} \right)^{\sqrt{5}} \right) = \sum_{k=0}^{\infty} \frac{1}{5^{5k}} \left(\frac{5}{5k+2} + \frac{1}{5k+3} \right)
$$

Some Base 10 BBP-Type Formulas

$$\begin{aligned}
\log\left(\frac{9}{10}\right) &= -\sum_{k=1}^{\infty} \frac{1}{k10^k} \\
\log\left(\frac{1111111111}{387420489}\right) &= 10^{-8} \sum_{k=0} \frac{1}{10^{10k}} \left(\frac{10^8}{10k+1} + \frac{10^7}{10k+2} + \ldots + \frac{1}{10k+9}\right) \\
\frac{\cos^{-1}(9/10)}{\sqrt{19}} &= \sum_{k=1}^{\infty} \frac{D_{k-1}}{k10^k},
\end{aligned}$$

where in the last line, D_k satisfy the recursion $D_0 = D_1 = 1,\ D_{k+1} = D_k - 5D_{k-1}$.

Is There a Base 10 BBP-Type Formula for Pi?

None is known. In fact, no BBP-type formula is known for π except in base 16 (which can be used to compute digits in any power-of-two base). In this sense 16 can be thought of as the "natural" base for π.

Do All BBP-Type Formulas Give Irrational Constants?

No. Examples:

$$
\begin{aligned}
1 &= \sum_{k=1}^{\infty} \left(\frac{2}{k} - \frac{1}{k+1} \right) \qquad \text{(a telescoping sum)} \\
0 &= \sum_{k=0}^{\infty} \frac{1}{16^k} \left(\frac{-8}{8k+1} + \frac{8}{8k+2} + \frac{4}{8k+3} + \frac{8}{8k+5} + \frac{2}{8k+6} - \frac{1}{8k+7} \right) \\
0 &= \sum_{k=0}^{\infty} \frac{1}{4096^k} \left(\frac{-256}{24k+5} + \frac{256}{24k+6} + \frac{128}{24k+7} + \frac{128}{24k+9} + \frac{-128}{24k+10} \right. \\
&\quad + \frac{-64}{24k+11} + \frac{-64}{24k+12} + \frac{-16}{24k+14} + \frac{24}{24k+16} + \frac{4}{24k+17} + \frac{-4}{24k+18} \\
&\quad \left. + \frac{-2}{24k+19} + \frac{-2}{24k+20} + \frac{-2}{24k+21} + \frac{-3}{24k+22} + \frac{1}{24k+23} \right)
\end{aligned}
$$

Further, by translating the indices of summation in any of these sums, an infinite class of *nonzero* sums can also be produced. Sequences corresponding to these formulas exhibit finite attractors, not equidistribution.

Some Basic Lemmas

The notation $\{x\}$ denotes the fractional part of x, i.e. x mod 1.

1. A number x is normal to base b iff the sequence $(\{b^d x\} \ : \ d = 1, 2, 3, \ldots)$ is equidistributed.
2. Assume x is normal to base b, and denote by r a nonzero rational number. Then rx is normal to base b; moreover x is also normal to any base $c = b^r$.
3. Assume a sequence (t_n) has the property that $t_n \to C$ as $n \to \infty$. Then a sequence $(\{x_n + t_n\})$ in $[0, 1)$ is equidistributed iff (x_n) is.
4. Assume a sequence (t_n) has the property that $t_n \to C$ as $n \to \infty$. Then a sequence $(\{x_n + t_n\})$ in $[0, 1)$ has a finite attractor iff (x_n) does.
5. Let α be real and $b \geq 2$ be an integer. If the sequence $x = (\{b^n \alpha\})$ has a finite attractor W, then W is a periodic attactor, and each $w_i \in W$ is rational.
6. If the sequence (x_n) as defined for Hypothesis A has a finite attractor W, then W is a periodic attractor, and each attractor point is rational.
7. The sequence $(\{b^n \alpha\})$ has a finite attractor iff α is rational.

Basic Theorem

Theorem 2. For a sequence $x = (x_n)$ as defined in Hypothesis A, define a real number α via a generalized polylogarithm series:

$$\alpha = \sum_{k=1}^{\infty} \frac{1}{b^k}\frac{p(k)}{q(k)}.$$

Then α is rational iff x has a finite (periodic) attractor.

Proof: From Lemma 6 we know that the sequence $(\{b^n\alpha\})$ has a periodic attractor iff α is rational. Following the BBP strategy, we can write

$$\begin{aligned} \{b^n\alpha\} &= \left(\sum_{k=1}^{n} \frac{b^{n-k}p(k)}{q(k)} + \sum_{k=n+1}^{\infty} \frac{b^{n-k}p(k)}{q(k)}\right) \bmod 1 \\ &= (x_n + t_n) \bmod 1 \end{aligned}$$

where x satisfies the recursion $x_0 = 0$, and

$$x_n = bx_{n-1} + \frac{p(n)}{q(n)},$$

Provided that $\deg p < \deg q$ as in Hypothesis A, we have $t_n \to 0$. Hence it follows from Lemma 4 that (x_n) has a periodic attractor iff α is rational.

Proof of Theorem 1

Theorem 1: Assuming Hypothesis A, each of the constants $\pi, \quad \log 2, \quad \zeta(3)$ is normal to base 2. Also, on Hypothesis A, if $\zeta(5)$ is irrational then it likewise is normal to base 2.

Proof. Each of the constants π, $\log 2$, $\zeta(3)$ is known to be irrational. Base 2 BBP-type formulas are known for each. Hence by Theorem 2, their associated sequences do not have periodic attractors. Thus, assuming Hypothesis A, their associated sequences are equidistributed, so that they are normal to base 2.

An Illustration of Theorem 1

Recall that

$$\log 2 = \sum_{k=1}^{\infty} \frac{1}{k2^k}$$

Let α_n be the binary expansion of $\log 2$ after n digits. Then we can write

$$\begin{aligned} \alpha_n &= \{2^n \log 2\} = \sum_{k=1}^{\infty} \frac{2^{n-k}}{k} \bmod 1 \\ &= \left(\sum_{k=1}^{n} \frac{2^{n-k} \bmod k}{k} \bmod 1 + \sum_{k=n+1}^{\infty} \frac{2^{n-k}}{k} \right) \bmod 1 \\ &= (x_n + t_n) \bmod 1 \end{aligned}$$

where $t_n \to 0$, and x_n satisfies the recursion $x_0 = 0$,

$$x_n = 2x_n + \frac{1}{n}$$

$\log 2$ is known to be irrational. Thus if Hypothesis A could be establish, it would follow that (x_n) is equidistributed, and that $\log 2$ is normal.

Can We Relax the Conditions of Hypothesis A?

1. Hypothesis A requires that $x_0 = 0$ (or at least some rational value). Consider the sequence associated with $\log 2$, namely

   $$x_n = 2x_{n-1} + \frac{1}{n} \bmod 1$$

   with $x_0 = 1 - \log 2$ instead of 0. The resulting sequence is *not* equidistributed — in fact, it converges to zero.

2. Hypothesis A requires that the perturbation term r_n be the quotient of two polynomials. Suppose we were to allow expressions such as $r_n = 1/2^{n^2-n}$. In this case the associated constant is

   $$\alpha = \sum_{n=1}^{\infty} \frac{1}{2^{n^2}}$$

   which is clearly irrational, but *not* normal to base 2.

A Curious Phenomenon in the Pi Iteration

Consider the binary sequence $y_k = \lfloor 2x_k \rfloor$, where x_k is the iteration for $\log 2$:

$$x_k = (2x_{k-1} + \frac{1}{k}) \bmod 1$$

The sequence (y_k) agrees well with the true binary digits of $\log 2$ — fifteen of the first 200 digits are incorrect, but only one in the range 5000 – 8000.

Now consider let $y_k = \lfloor 16x_k \rfloor$, where x_k is the iteration for π:

$$x_k = \left(16x_{k-1} + \frac{120k^2 - 89k + 16}{512k^4 - 1024k^3 + 712k^2 - 206k + 21}\right) \bmod 1$$

In this case, the hex sequence (y_k) appears to agree *exactly* with the true hex digits of π — there are no errors in the first 100,000 digits.

A Connection to Pseudorandom Number Generators

Consider the canonical case $\alpha = \log 2$. One can write

$$\{2^d \alpha\} = \left(\frac{2^{d-1} \bmod 1}{1} + \frac{2^{d-2} \bmod 2}{2} + \cdots + \frac{1}{d} + t_d \right) \bmod 1$$

Now fix an integer D, and consider this iteration:

$$R(D, k) = \left(\frac{2^{k-1} \bmod 1}{1} + \frac{2^{k-2} \bmod 2}{2} + + \cdots + \frac{2^{k-D} \bmod D}{D} \right) \bmod 1.$$

As k advances, this is a sum of normalized linear congruential pseudorandom number generators.

Question: What is the period of this type of "cascaded" pseudorandom number generator? Empirical studies suggest it increases exponentially with D, but we have no rigorous results. More research is needed here.

Open Questions

- Is there a natural generalization perturbation function r_n in Hypothesis A?
- Can we apply more of the theory of ergodic systems and chaotic-dynamic maps to these questions?
- Can we develop a more complete theory of the special instances in which a generalized polylogarithm series has a rational sum?
- Can we make more inroads into the theory of cascaded linear congruential pseudorandom number generators?
- Can we obtain formal bounds on the lengths of periods produced by cascaded pseudorandom number generators?
- Can we deal with algebraic irrationals (such as $\sqrt{2}$ and the golden mean τ) in this theory?

For Full Details

See the manuscript "On the Random Character of Fundamental Constant Expansions", which is available from either of our web sites

`http://www.nersc.gov/~dhbailey`
`http://www.perfsci.com`

A second paper, "Random Generators and Normal Numbers", will be available soon from these same web sites.